Interoperability and Advanced Computational Technologies in the GSFC Land Information System

Project-leads: Christa Peters-Lidard & Paul Houser, Code 974

GSFC Team: Y. Tian, S. V. Kumar, J. Geiger, S. Olden,
L. Lighty, J. L. Eastman

Princeton Team: J. Sheffield , E. F. Wood

NOAA Team: K. Mitchell , H. Wei, G. Gayno

COLA Team: P. Dirmeyer, B. Doty and J. Adams

Land Information System (LIS)
http://lis.gsfc.nasa.gov

- Definition
  - LIS is a high performance land surface modeling and data assimilation system, based on GSFC's Land Data Assimilation Systems (http://ldas/gsfc.nasa.gov).
- Acknowledgement
  - LIS is a Grand Challenge Investigation funded under NASA ESTO/CT CAN-00-OES-01, with additional support from NASA ESTO/AIST NRA-02-OES-04.
- Outline of Talk
  - High-resolution land surface modeling
  - Interoperability (ESMF)
  - High-performance computing and communication

LIS Modeling Approach

Inputs

Topography, Soils (Static)

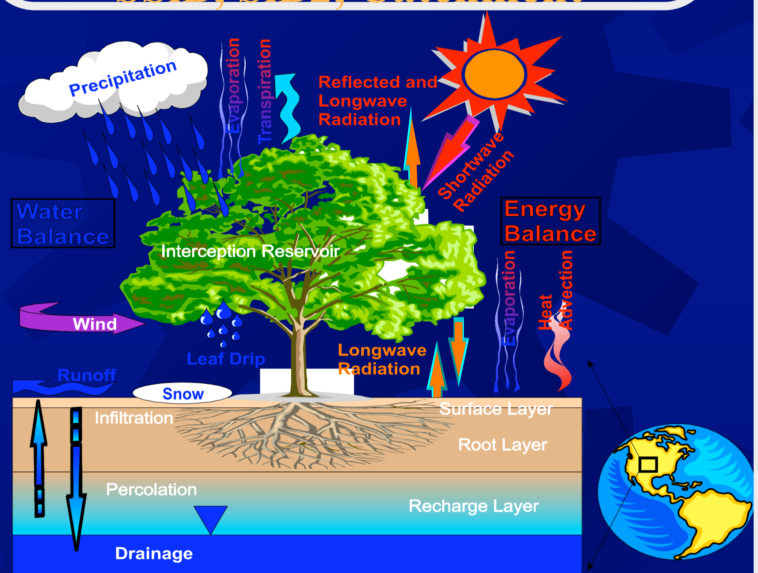
MODIS/AVHRR Land Cover, Leaf Area Index (Monthly)

Modeled + TRMM/SSM/I Meteorology (Hourly-3 hourly)

Observed Surface States (e.g., Snow, Soil Moisture)

Physics

Land Surface Models (LSM)
(Time steps=min-hr
Spatial grid=m-deg)
Noah, CLM, VIC, Mosaic, SSiB, SiB2, Catchment



Data Assimilation Modules (EnKF, EKF)
Physical Space Analysis System (PSAS) 3-D VAR
Rule-based

Outputs

Soil Moisture & Temperature Profiles

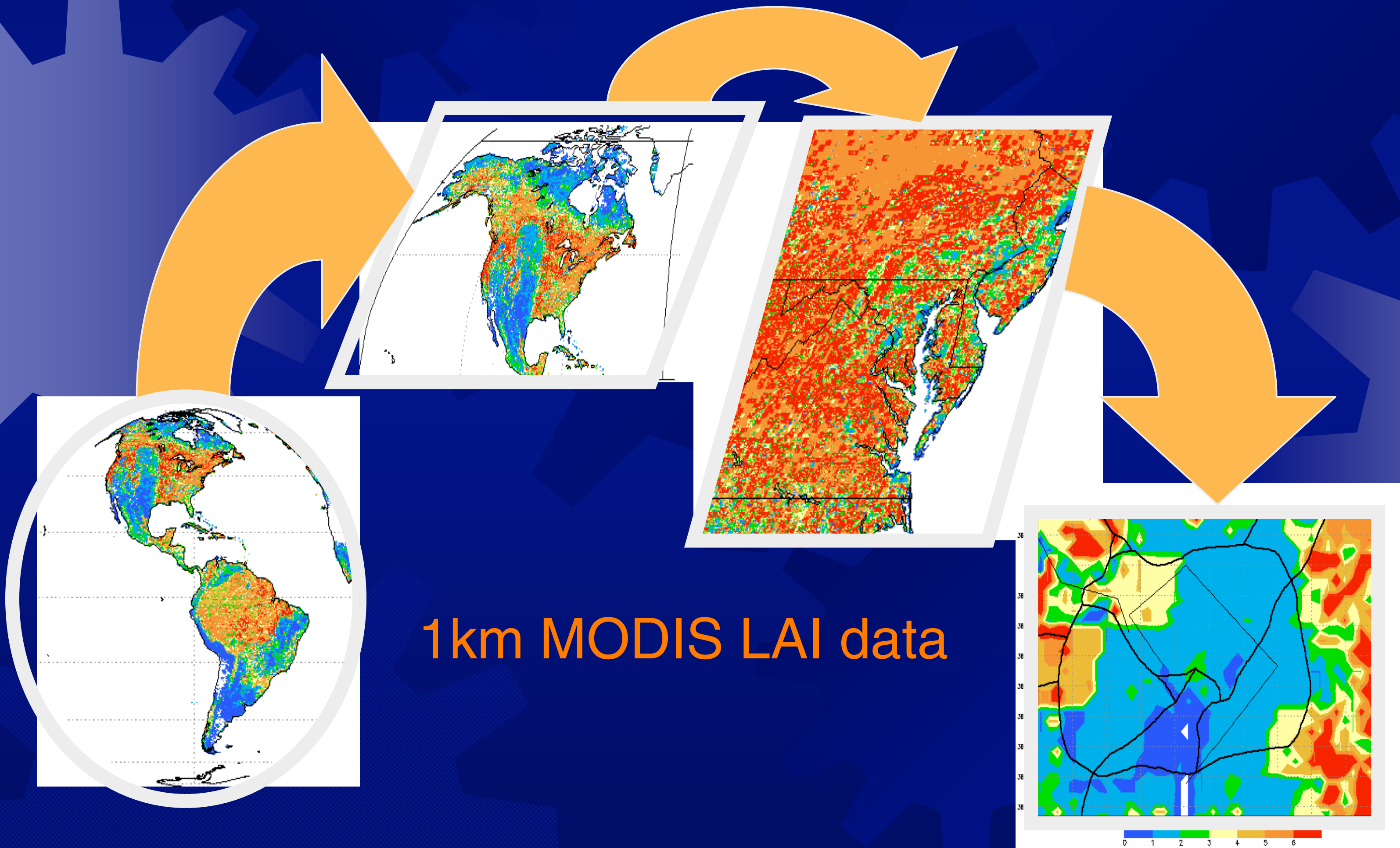
Surface Energy Fluxes (e.g., H,LE) → **Atm. Models (GCE/WRF)**

Surface Water Fluxes (e.g.,Runoff) → **Water Resources/ Ocean Models**

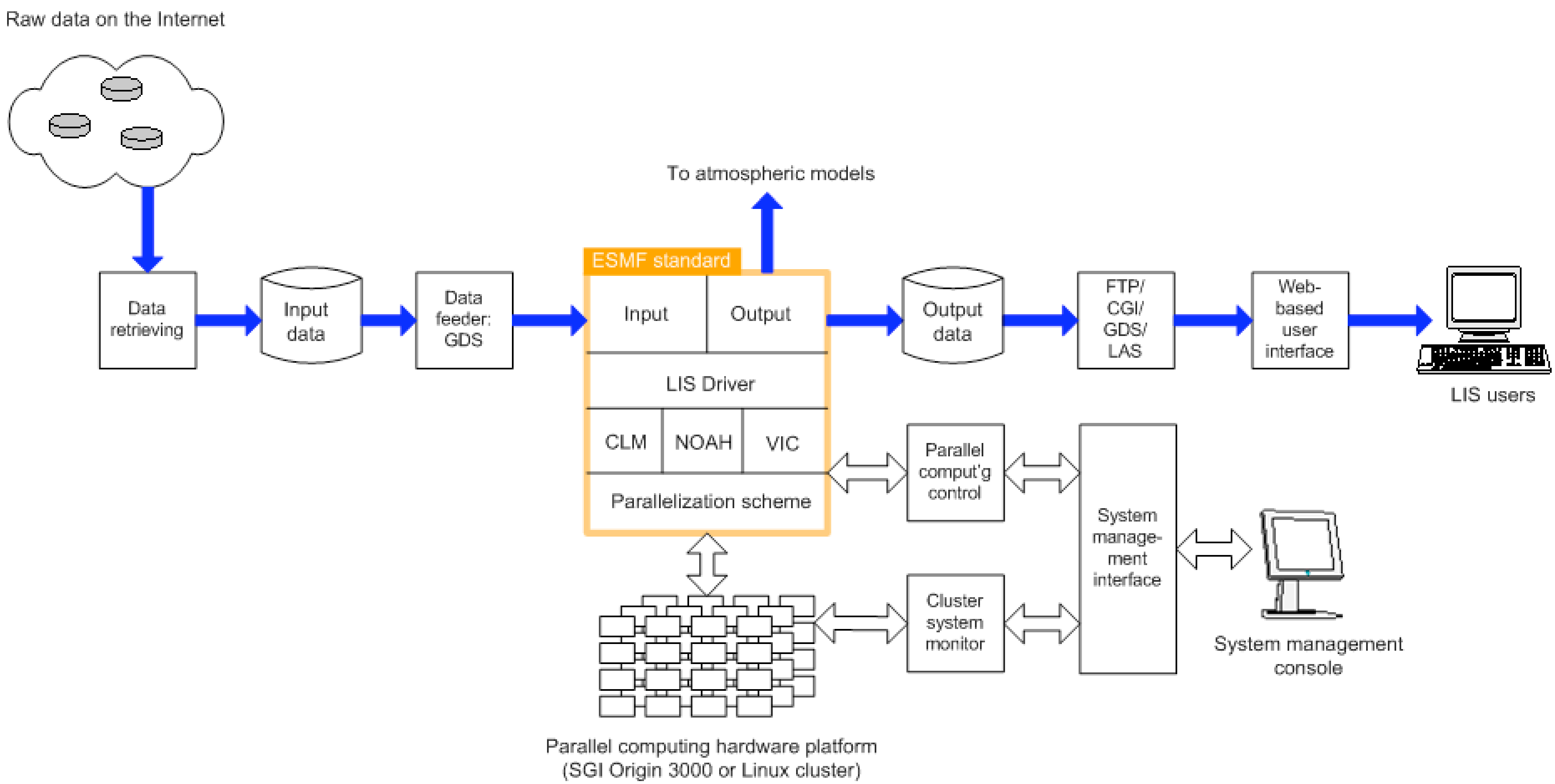
Surface States: Snowpack LAI (some)

Example: MODIS Leaf Area Index Product in LIS

Key LIS Objective: To model globally at the same resolution as EOS data



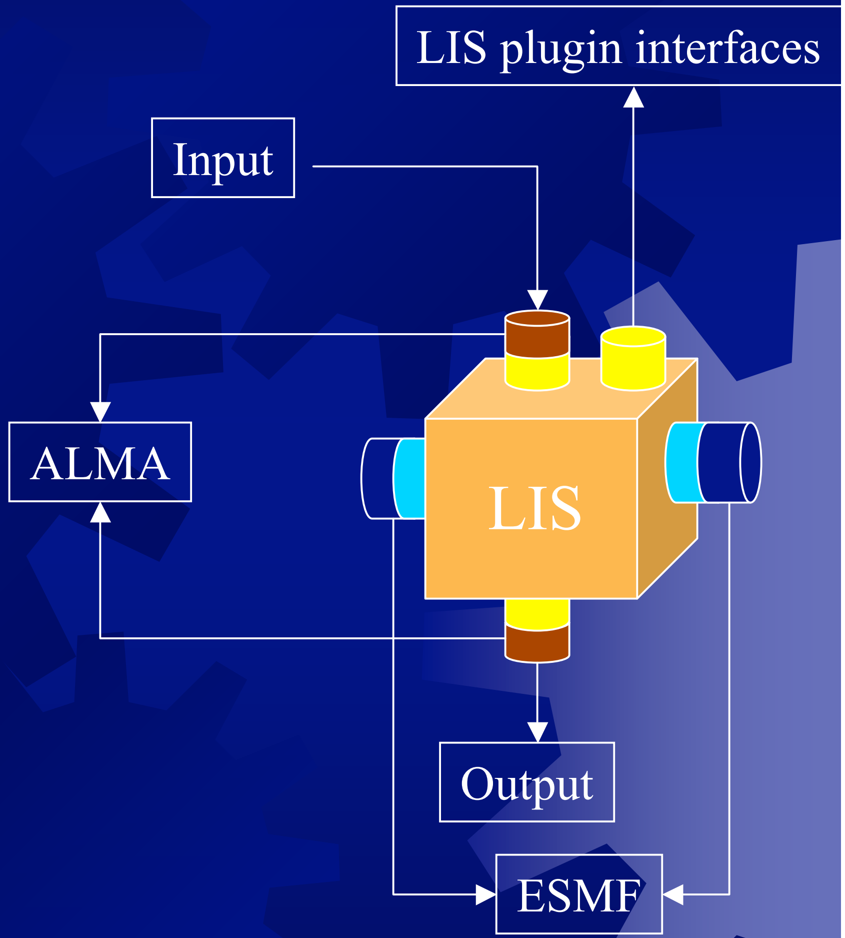
Overview of LIS Architecture



Interoperability in LIS – ESMF and ALMA

- LIS adopts the utilities and compliance guidelines provided by the Earth System Modeling Framework (ESMF)

- LIS adopts the variable naming, units and sign conventions of the international Assistance for Land Modeling Activities (ALMA) standards



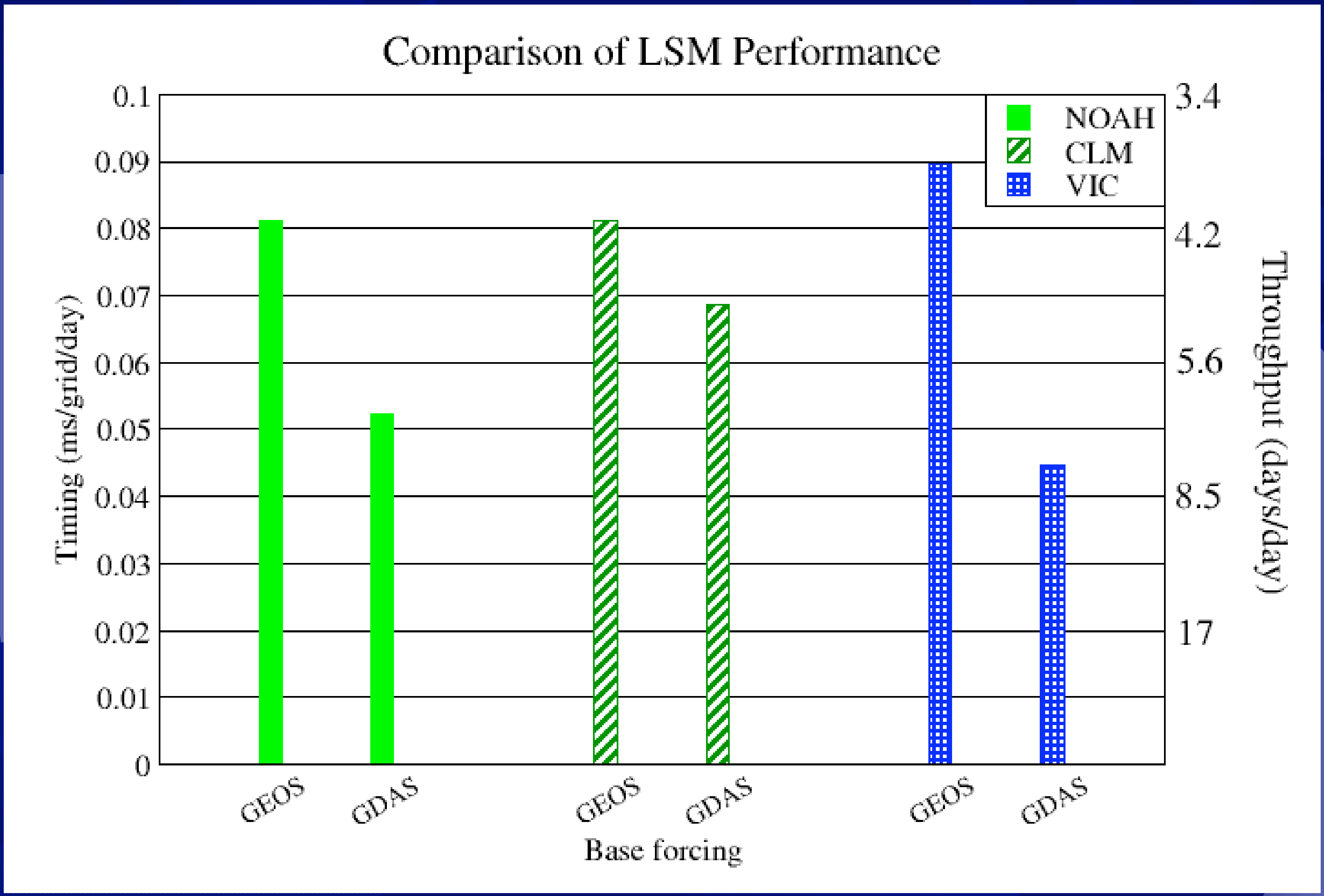
Computational Challenges:

Resolution	1/4 deg	5 km	1 km
Land Grid Points	2.43E+05	5.73E+06	1.44E+08
Disk Space/Day (Gb)	1	28	694
Memory (Gb)	3	62	1561

Computational Goals:

- High-resolution global land surface simulation –1km
- Real-time simulation – 1 day in a day
- Running semi-operational
  - High throughput
  - Custom commodity cluster
  - Scalable
  - Reliable

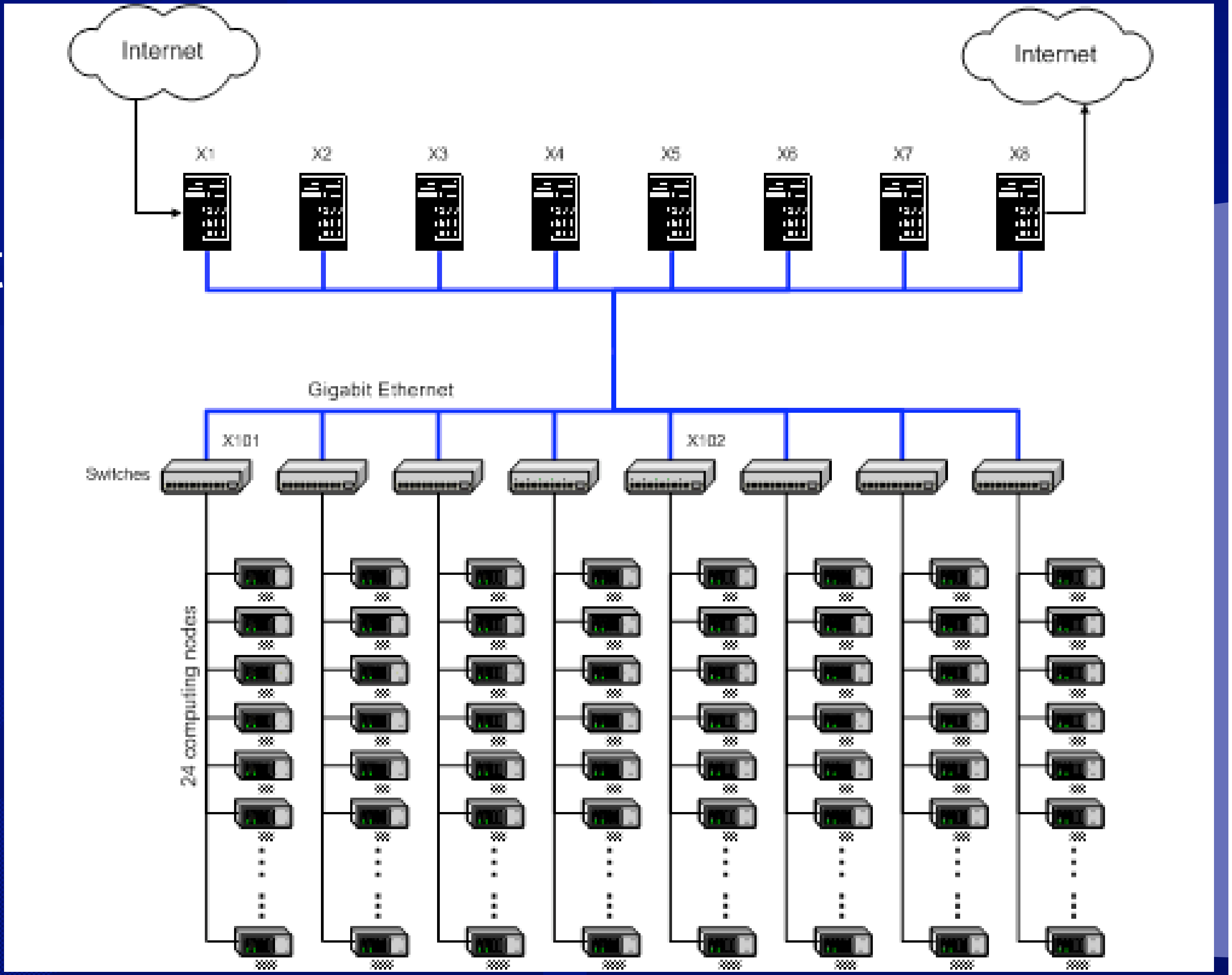
Milestone G Achieved: 1 km, >3 days/day



Compute nodes used: 128

LIS’ custom commodity cluster

- Linux
- Custom job management system
- Strong fault-tolerance
- Parallel IO support

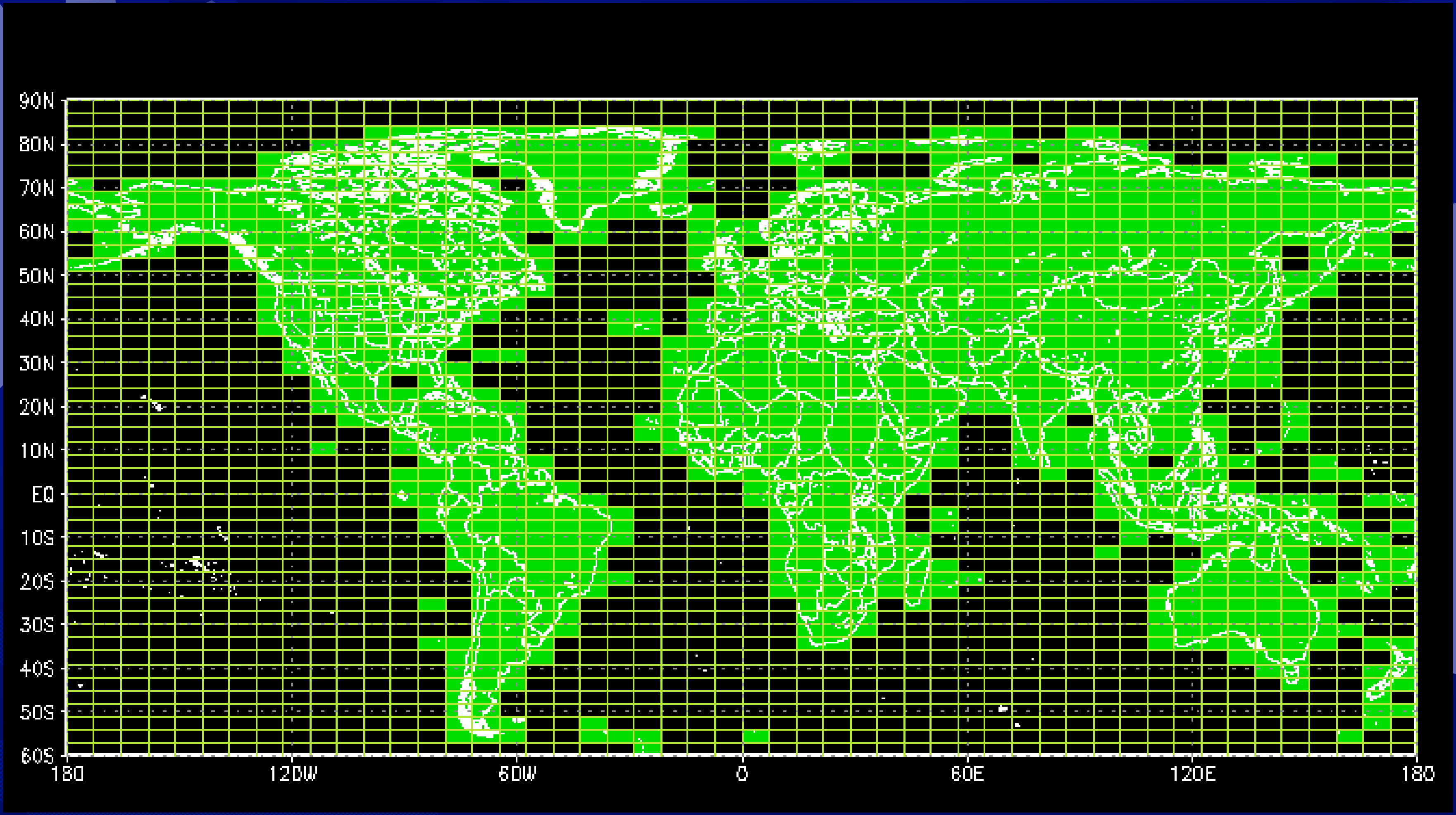


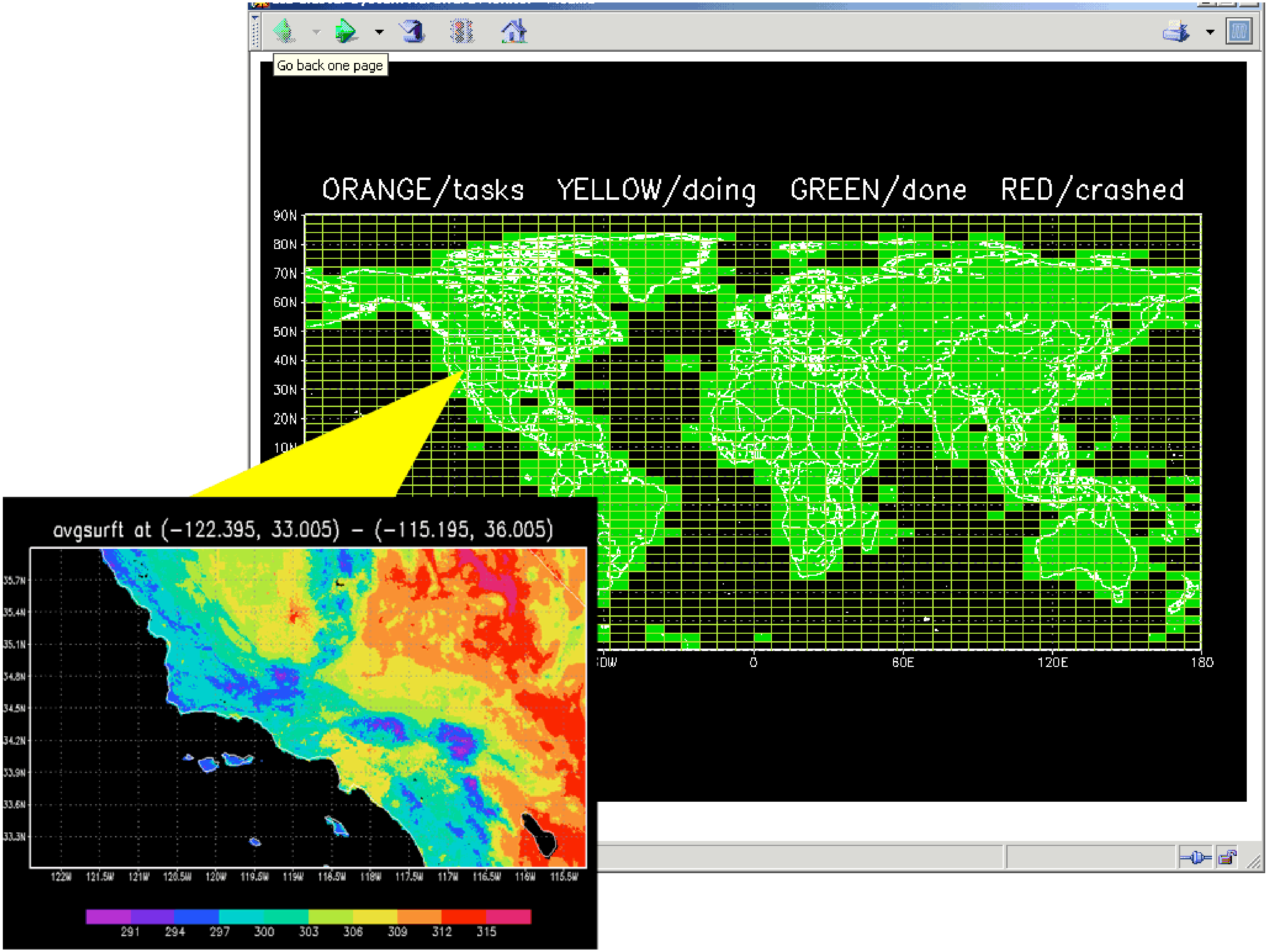
LIS’ job management system
-- A “pool of tasks” scheme:

- Exploits the coarse-grained feature
- De-synchronize IO and data access
- Optimal resource utilization
- Strong fault-tolerance
- Real-time visualization of progress and results

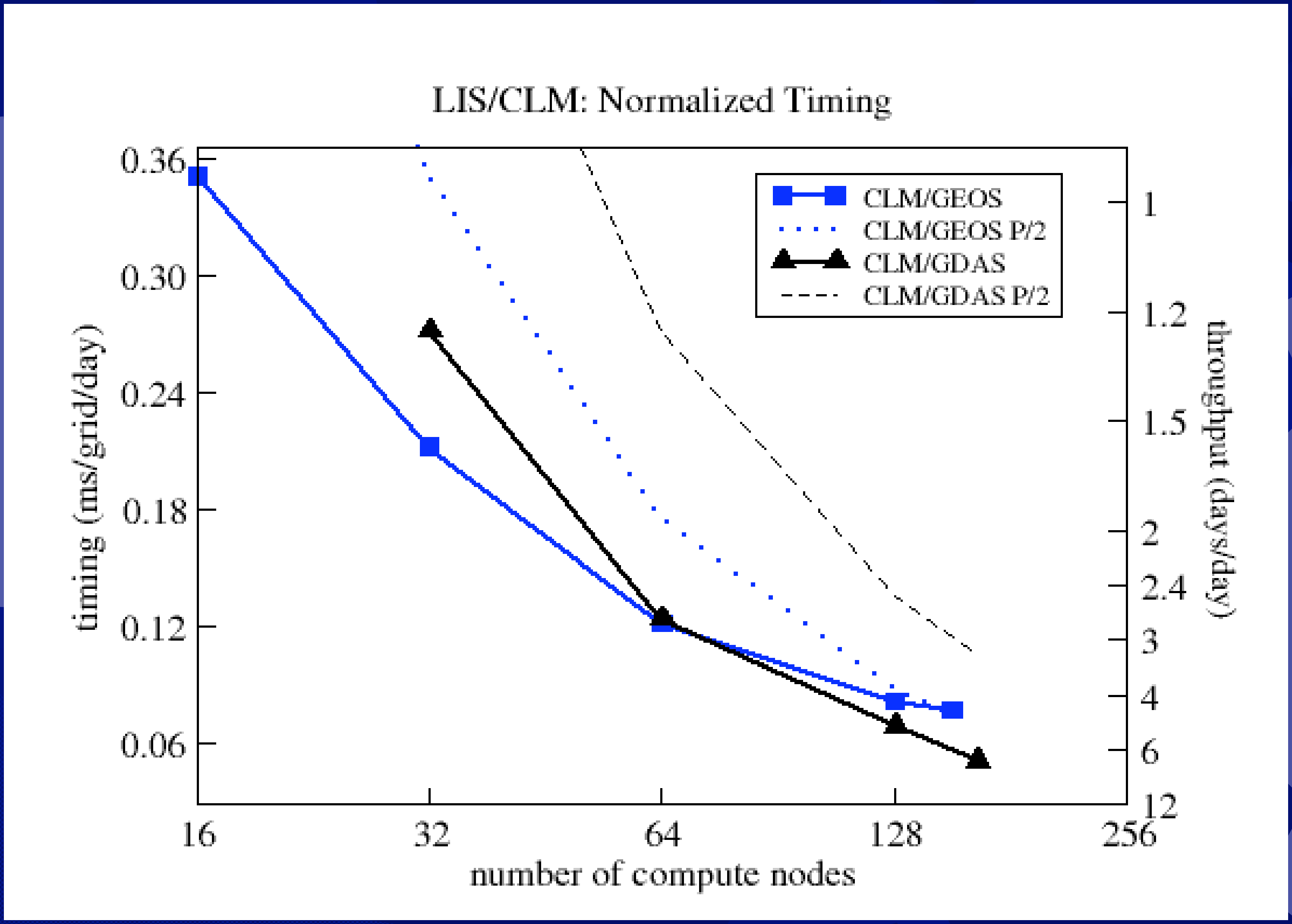
LIS' coarse-grained parallelism

- No communication between sub-domains
- De-synchronize IO and data access to eliminate bottleneck
- Linear scalability in theory




Go back one page
ORANGE/tasks YELLOW/doing GREEN/done RED/crashed
90N
80N
70N
60N
50N
40N
30N
20N
0
60E
120E
180
avgsurft at (−122.395, 33.005) − (−115.195, 36.005)
291 294 297 300 303 306 309 312 315

How LIS scales:



Resolution	1/4 deg	5 km	1 km
Land Grid Points	2.43E+05	5.73E+06	1.44E+08
Disk Space/Day (Gb)	1	28	694
Memory (Gb)	3	62	1561

LIS' Biggest challenge: IO

- Earth science data volume grows exponentially
- CPU performance follows Moore's Law
- Disk capacity also follows Moore's Law
- However, disk IO speed lagged far behind (~10% improvement per year)

High-performance IO is critical to LIS, as well as to other Earth modeling systems

Overcoming the IO Bottleneck

- LIS' innovations
  - On-demand data serving with GrADS-DODS servers (GDS)
  - Parallel input data serving with load balancing
  - Distributed output and storage
  - Serving users off distributed storage
  - Highly scalable data replication with peer-to-peer (P2P) technology

GrADS-DODS: Grid Analysis and Display System / Distributed Oceanographic Data System

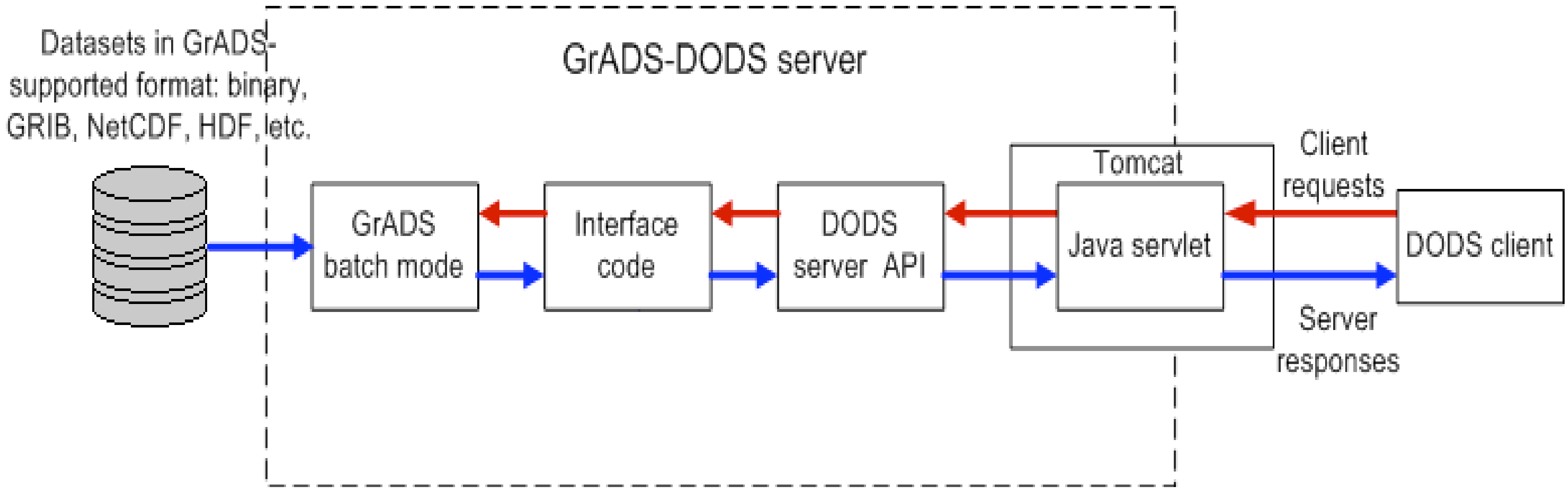
GrADS-DODS Server (GDS)

http://grads.iges.org/grads/gds/

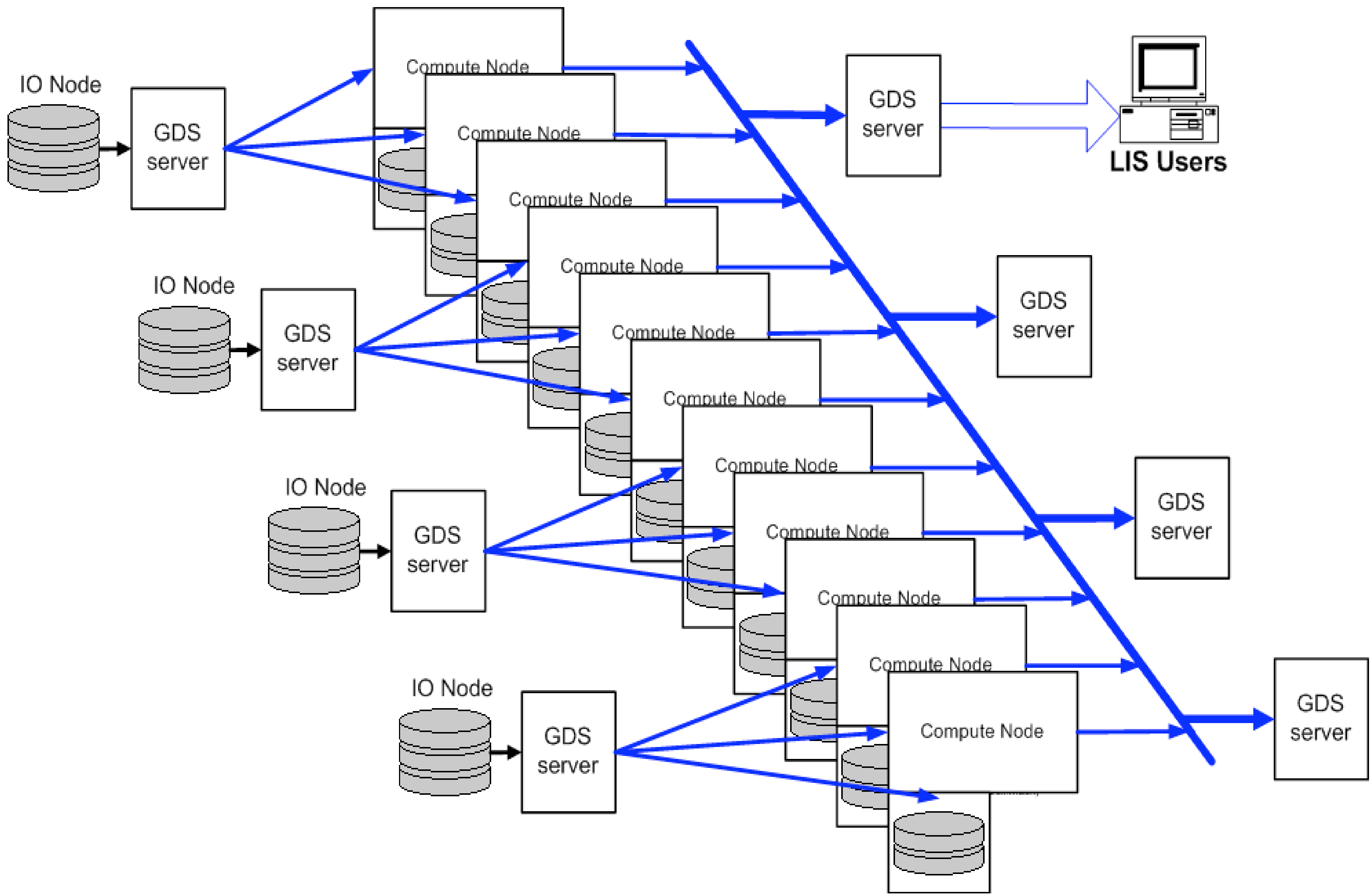
- **Share data:** Enterprise-wide; Internet-wide. Data-format independent
- **Data interoperability:** Consistent metadata for many data types
- **Distributed analysis:** Reduces network load; improves interactivity
- **On-demand data serving:** Users can request arbitrary subsets of data on the fly (no need to download files via ftp/http, etc)

GrADS-DODS Server (GDS)

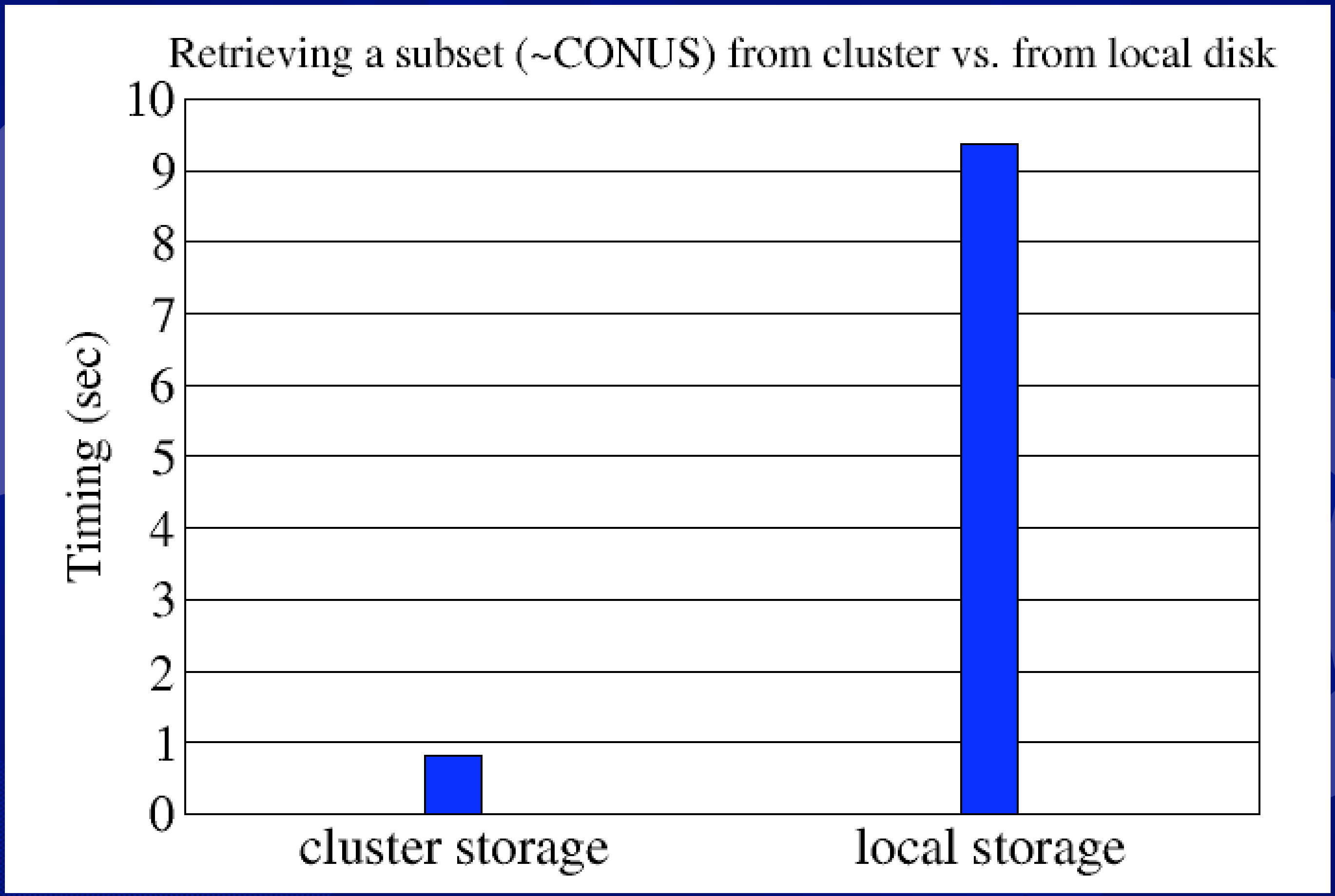
Client-Server Architecture



LIS parallel IO architecture

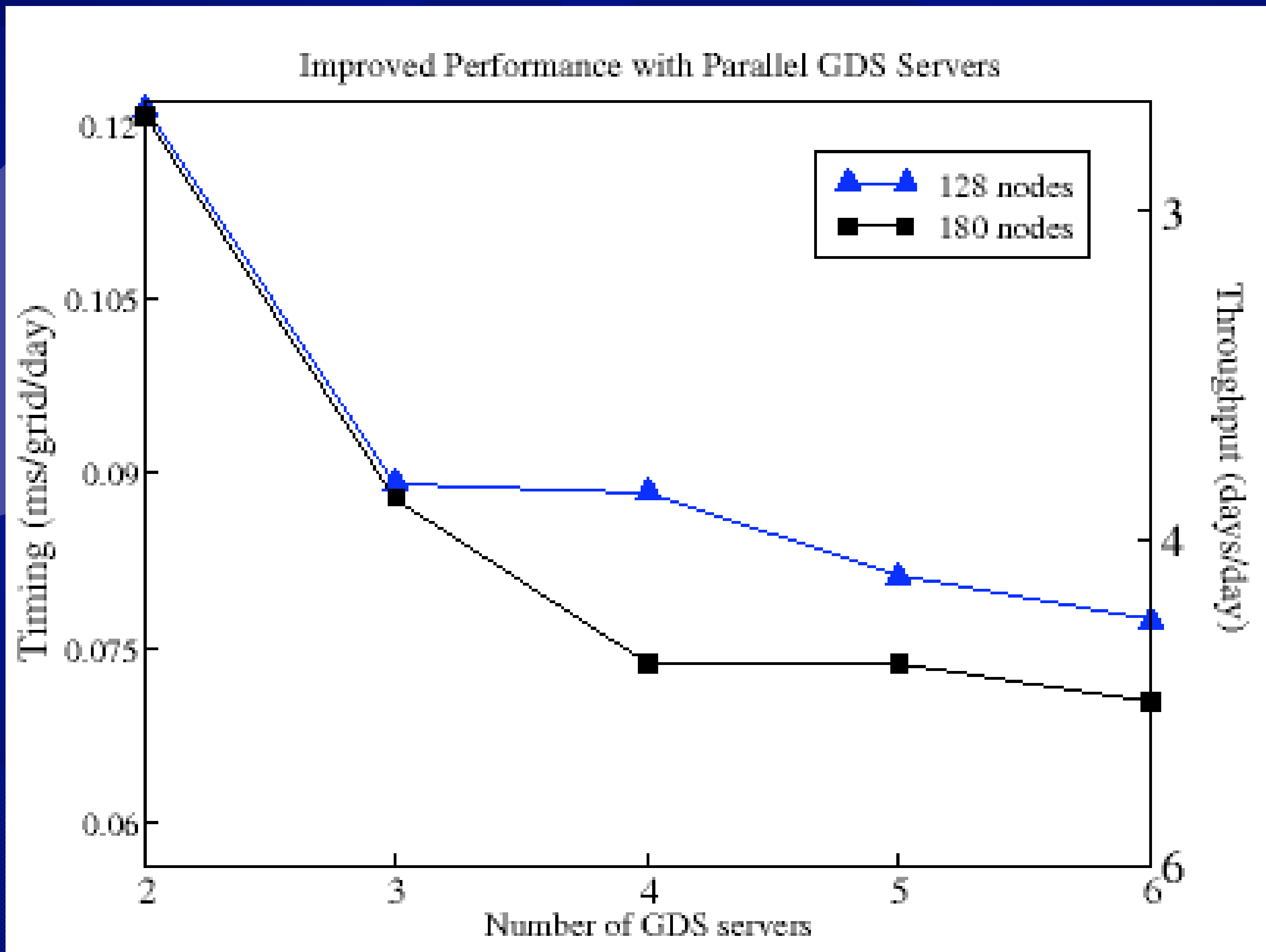


Performance improvement from LIS distributed storage



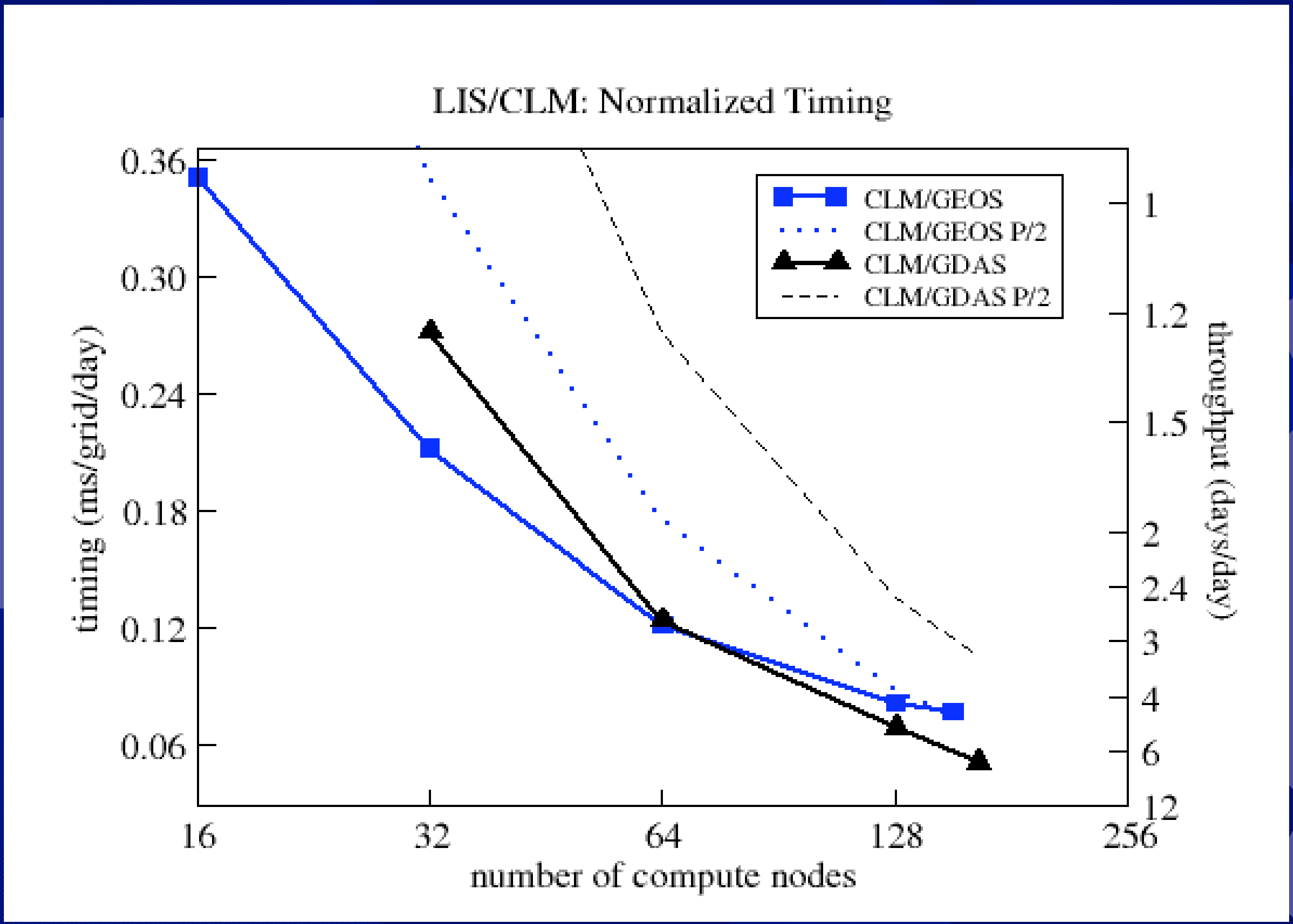
CONUS=CONtinental U.S.

Performance vs. Parallel GDS IO



Performance vs. Processors

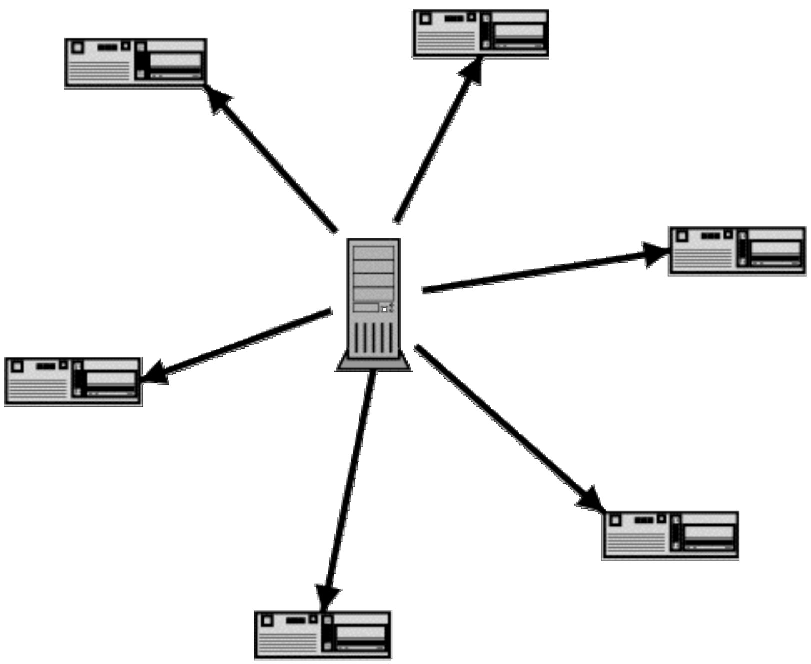
5 GDS and parallel IO



Application of peer-to-peer technology on Linux cluster

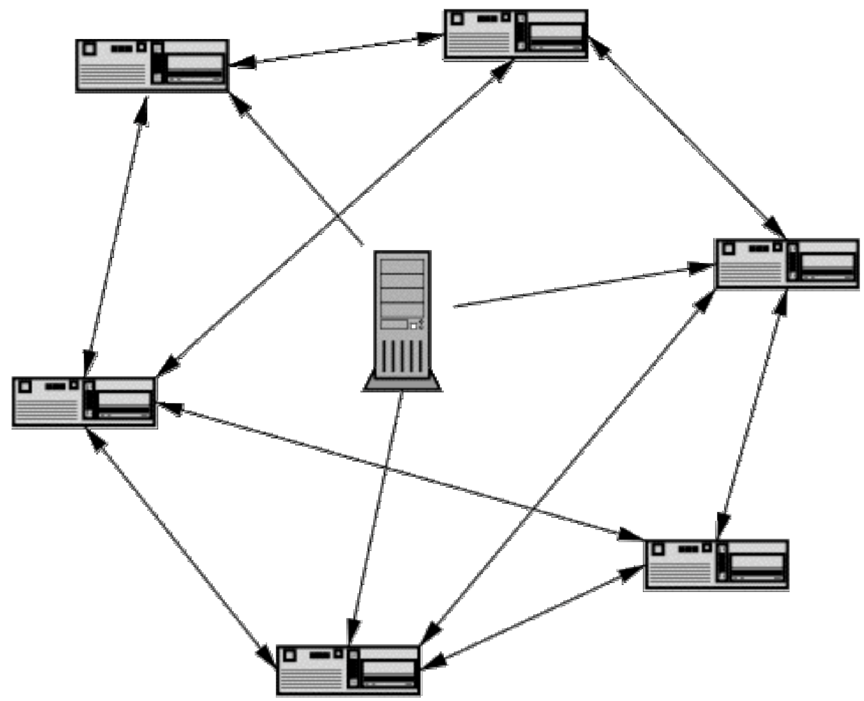
- Conventional client-server system does not work well with large data and many clients
- LIS needs to get copies of data (e.g., static parameters, OS images) on ~200 nodes quickly and reliably
- P2P is not only for MP3 sharing. We use it for data sharing on the cluster
- Much higher performance and better scalability than NFS.

Client-server architecture



Peer-to-peer architecture

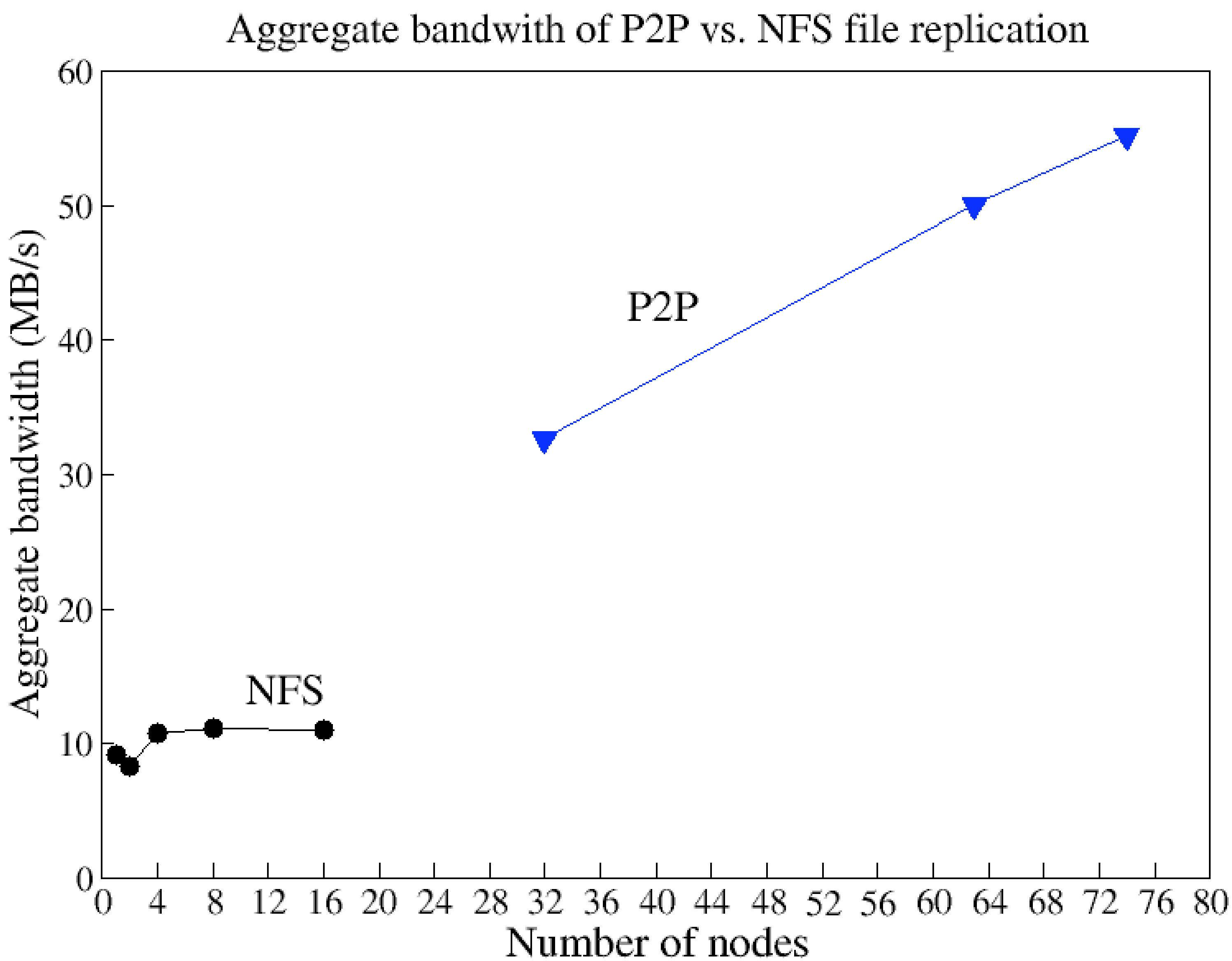
http://bitconjurer.org/BitTorrent/



Advantages of P2P over client-server:

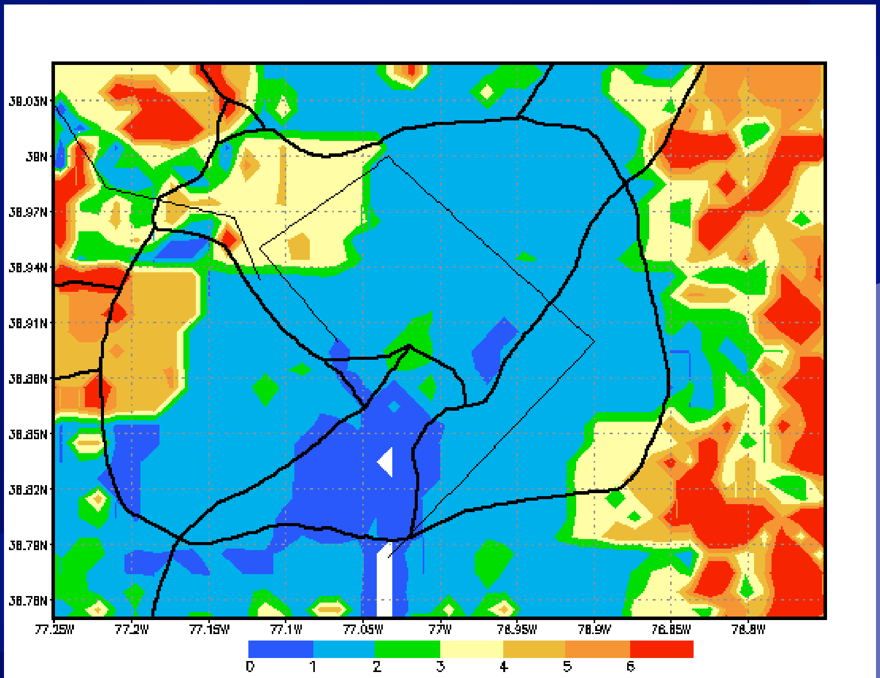
- Highly scalable: better performance with more peers
- Reliable: guaranteed data integrity
- Easy to manage: no scheduling, auto-resumption of transfer

Replicating 2GB data on the nodes:



Summary

- LIS is exploring technologies for next generation Earth system modeling
  - Global simulation on low-cost parallel computing platform with fault-tolerance
  - Breaking the IO barrier with parallel and distributed technologies, and on-demand data serving
  - ESMF compliance allows other Earth models to couple with LIS and tap LIS' data and computing power

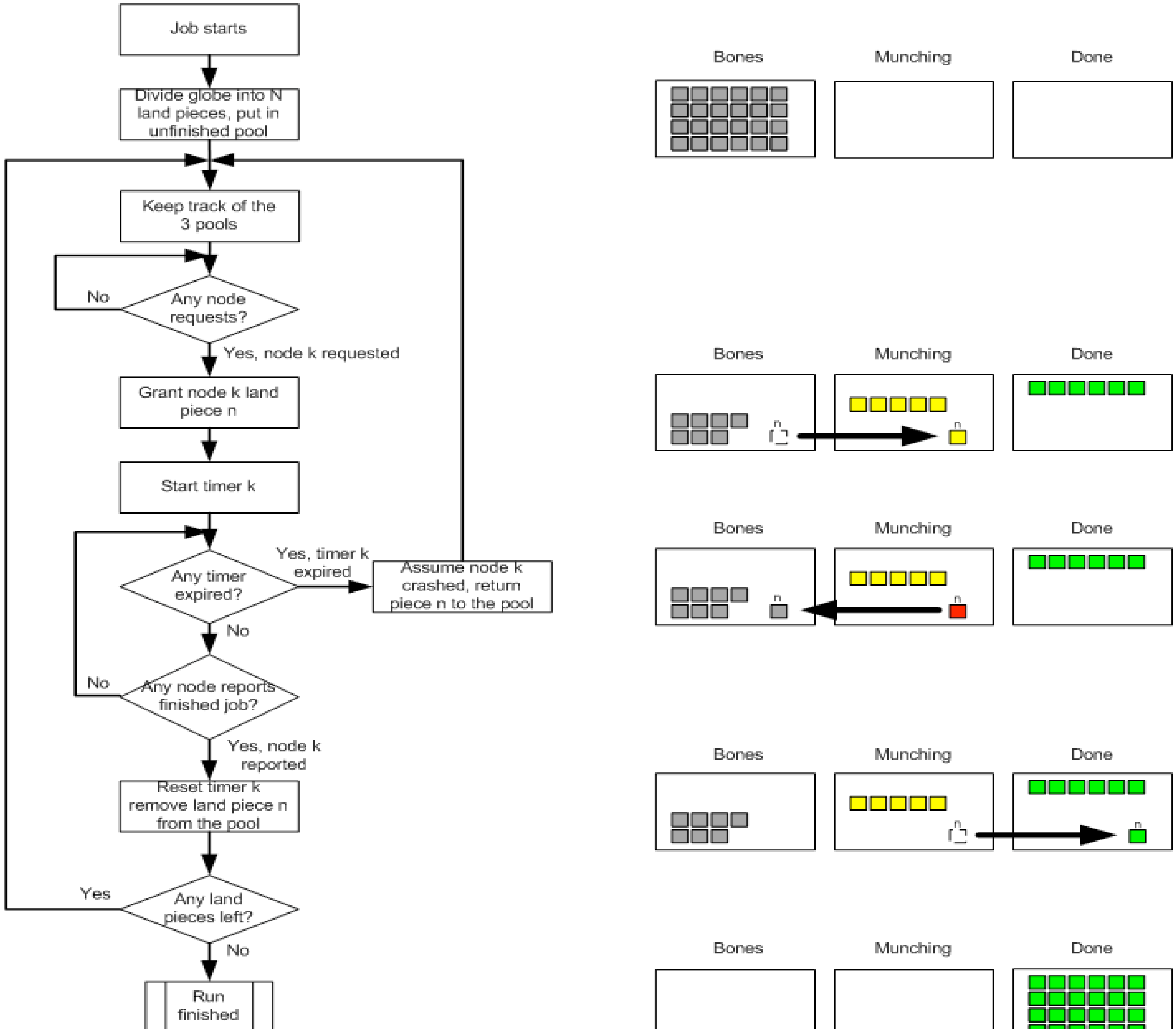


- LIS is pushing the envelope of Earth system modeling

LIS Extensions

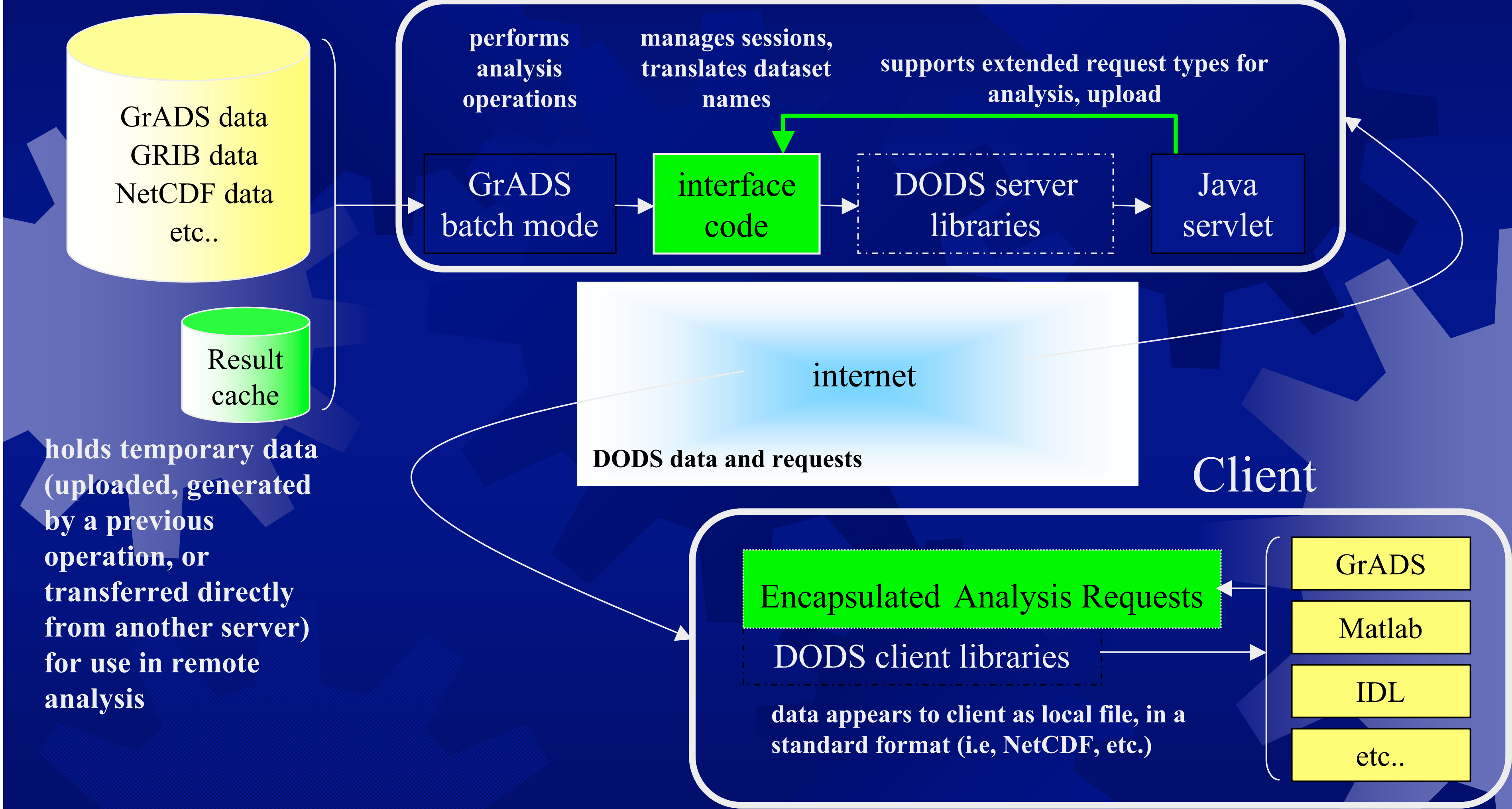
1. Coupling High-Resolution Earth System Models (AIST02 NRA)

2. The Impact of Precipitation Measurement Missions on Water Resources (Precipitation02 NRA)


Job starts
Divide globe into N land pieces, put in unfinished pool
Keep track of the 3 pools
Any node requests?
No
Yes, node k requested
Grant node k land piece n
Start timer k
Any timer expired?
Yes, timer k expired
Assume node k crashed, return piece n to the pool
No
Any node reports finished job?
No
Yes, node k reported
Reset timer k remove land piece n from the pool
Any land pieces left?
Yes
No
Run finished
Bones
Munching
Done
n
Bones
Munching
Done
n
Bones
Munching
Done
n
Bones
Munching
Done
n
Bones
Munching
Done

GrADS-DODS Server (GDS)

Client-Server Architecture



Joe Wielgosz: 5/25/00